JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

J. Agric. Food Chem. **2006,** *54,* 6951−6956 **6951**

Measurement of Inorganic Phosphorus in Soybeans with Near-Infrared Spectroscopy

STEPHEN R. DELWICHE,*,† LESTER O. PORDESIMO,†,‡ ANDREW M. SCABOO,§,# AND VINCENT R. PANTALONE§

Beltsville Agricultural Research Center, Instrumentation and Sensing Laboratory, Building 303, BARC-East, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, Maryland 20705-2350, and Department of Plant Sciences, University of Tennessee, Knoxville, Tennessee 37996

This study explored the feasibility of near-infrared (NIR) quantitative and qualitative models for soybean inorganic phosphorus (P_i), which is complementary to phytic acid, a component of nutritional and environmental importance. Spectra, consisting of diffuse reflectance (1100−2500 nm) of ground meal and single-bean transmittance (600−1900 nm) of whole seed, were collected on 191 recombinant inbred soybean lines. Partial least-squares regression models were individually developed for soy meal diffuse reflectance, single-bean transmittance, and averaged (24 beans/line) whole seed transmittance data. The best performance was obtained with diffuse reflectance data, in which the standard errors (rmsd) were 263 and 248 mg/kg for cross-validation and validation sets, respectively. Model accuracy was lower for the 24-bean average transmittance spectra and still lower for single beans. Despite the overall poorer modeling ability of P_i with respect to the common macronutrient NIR regressions, such as those for protein and oil, this technique holds promise for use in breeding programs.

KEYWORDS: Soybean; phosphorus; near-infrared; single seed; breeding; phytate

INTRODUCTION

Phytic acid (*myo*-inositol-1,2,3,4,5,6-hexakisphosphate) is the primary storage compound for phosphorus in soybean [*Glycine max* (L.) Merr.] and other seeds, comprising approximately three-fourths of the total P content, with inorganic P (P_i) comprising slightly less than 25%. Inorganic P has a very strong negative correlation to phytate (*1*, *2*), such that the total level of P in soybean remains nearly constant. Theorized to be a storage mechanism for seed phosphorus (*1*), phytic acid is a cause of concern from nutritional and environmental standpoints. When consumed by humans and nonruminant animals such as poultry and swine, the seed's phytic acid can act as a chelating agent of minerals such as calcium, iron, zinc, and other cations, thereby preventing the absorption of such essential minerals by the animal (*3*). Feed rations are often supplemented with P_i or with the enzyme phytase that allows the capture of P from phytate. Phosphorus-laden poultry and swine manure is considered a significant source of water pollution (*4*).

Through recent breeding developments, nonlethal recessive mutations that result in soybean genotypes of low phytic acid have been developed (*5*, *6*). Oltmans et al. (*7*) demonstrated that the condition of low phytate is recessive and controlled by two alleles, designated *pha1* and *pha2*. The low-phytate condition occurs when both recessive alleles are homozygous. Most recently, Walker et al. (*8*) identified two SSR marker loci (Satt237 and Satt561) positively correlated with the quantitative trait loci (QTL) for soybean seed phytate. They found that 41% of the observed variation in inorganic phosphorus was attributed to the first linkage group, with the second linkage group accounting for 11% of the variation and both collectively accounting for an additional 8−11%. Although such markers hold promise for use in breeding programs for the selection of low phytate, phenotypic selection is still found to be at least as effective and, at this time, a more cost efficient method for the selection of the low-phytate condition (*9*). Therefore, a rapid phenotypic screening method that is capable of measuring soy phytic acid or P_i concentration and thereby performing a high versus low classification is of immediate value. Whether by cutoff value of concentration or by a direct classification algorithm, high versus low P_i categorization is based on the phenomenon that mature wild-type seed from cereals and oilseeds typically have P_i concentrations of <500 mg/kg, whereas the concentration in low-phytate lines will often exceed 1000 mg/kg (*10*).

Breeding programs are currently active in developing new cultivars with low phytate. In practice, screening for *pha1* and *pha2* is accomplished by examination of the phenotypic

* Author to whom correspondence should be addressed [e-mail delwiche@ba.ars.usda.gov; telephone (301) 504-8450].
† U.S. Department of Agriculture.
‡ Present address: Eastern Regional Research Center, USDA-ARS, Wyndmoor, PA.
§ University of Tennessee.
# Present address: Crop, Soil, and Environmental Sciences, University of Arkansas, Fayetteville, AR.

10.1021/jf060919n CCC: $33.50
Published on Web 08/29/2006

expression and most often by seed P_i assay, which is much less time-consuming than phytate analysis. The successful establishment of a near-infrared (NIR) spectroscopy-based measurement procedure for phytate or P_i would hold promise as a very useful tool in screening of early generations. Because the current study is a component of an ongoing low-phytate soybean breeding program in which lines are routinely screened and selected for high P_i (*5*), NIR calibration development is focused on this form of phosphorus. Ideally, NIR-based screening would be applied at the level of the individual bean, preferably by nondestructive means. Therefore, the objective of this study was to evaluate the potential of NIR reflectance of soybean meal and NIR transmittance of individual beans, with partial least-squares regression (PLSR) quantitative analysis and principal component linear discriminant analysis, for the measurement and classification of P_i.

MATERIALS AND METHODS

Plant Materials. Soybean lines arose from the initial cross of a high-yielding commercial cultivar with a low-phytate experimental line. Specifically, cultivar '5601T' (maturity group V) (*11*) was crossed with a low-phytate mutant germplasm Cx1834-1-2 (maturity group III, developed by J. R. Wilcox, USDA-ARS, West Lafayette, IN) (*5*). The initial cross was made in the summer of 2000, with F_1 seeds advanced to the F_5 stage in Costa Rica by single seed descent (*12*). Single plants ($F_{4:5}$) were grown in Knoxville, TN, in 2002 for the development of a recombinant inbred line (RIL) population. Approximately 700 F_5 single plants were harvested and sent to Costa Rica for seed increase and then sent back to Knoxville for planting in 3-m progeny rows. From 187 randomly selected individual progeny rows, three maturity groups, based on observations, were planted in two row plots (6.1 m per row) in a randomized complete block design in 2004. On the basis of soil test results, fertilizer for phosphorus and potassium (0−20−20) was applied to the plots before planting according to University of Tennessee Extension Service recommendation levels. Samples used in this study were from one field replicate grown at the East Tennessee Research and Education Center in Knoxville, TN. The soil phosphorus contents, determined from the average of 12 random soil samples within the replicated block, were 61, 59, and 19 mg/kg for the three maturity groups, corresponding to average seed P_i concentrations of 757, 931, and 819 mg/kg, respectively. With parents and checks, a total of 191 samples were available for NIR analyses, which was considered to be sufficient for evaluating the NIR method.

The authors recognize the importance that environment plays on seed characteristics and understand that genotype × environment interaction is an important consideration in selection strategies. The goal of this work was not to characterize soybeans grown under diverse environments, but rather to minimize environmental noise by focusing on seeds produced in a single uniform field environment from a genetic population known to exhibit a wide range in P_i. This strategy enabled us to evaluate the effectiveness of NIR for P_i prediction.

Equipment. One spectrometer, a model 6500 scanning monochromator (Foss-NIRSystems, Silver Spring, MD) with Vision software (version 2.50), was used to collect single-bean transmittance and soybean meal reflectance spectra. For single-bean transmittance, the monochromator was configured with a single tablet analyzer attachment. Light was conveyed from the exit slit of the monochromator through a fiber optic assembly with its 4-mm-diameter terminus positioned in direct overhead contact with the soybean (**Figure 1**). A standard type PIN indium gallium arsenide detector (3 mm × 3 mm) was positioned ≈2 mm directly below the soybean. A machined aluminum cartridge possessing an 8.5-mm-diameter by 4-mm-deep well, with a concentric 6.5-mm diameter (0.5-mm rim thickness) aperture at the well base, was used to house the soybean. Gaps between the bean and the circular aperture, caused by the imperfect radial symmetry of the bean, were filled with black moldable filler material (Silly Putty) to minimize stray light from striking the detector. A similar cartridge possessing two stacked panels of 1.5-mm-thick white Teflon sheet was used as an optical reference. For diffuse reflectance of soybean meal, the same monochromator was used, but with a spinning sample cup attachment in lieu of the tablet analyzer.

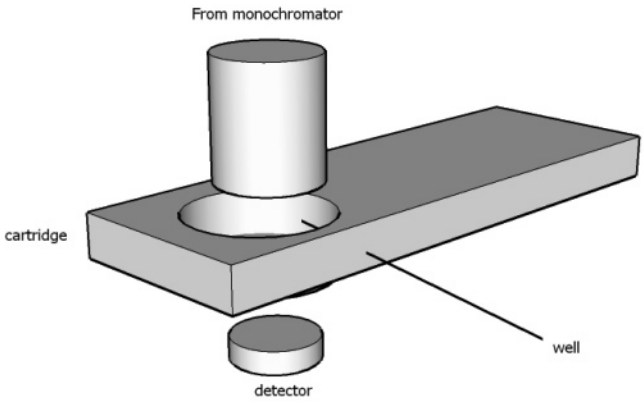

Figure 1. Schematic of single-bean transmittance apparatus.

Procedure. Single-bean transmittance spectra were collected on 24 randomly drawn soybeans from each sample. Each soybean was placed on its side in the cartridge, with the hilum of the bean at a constant radial position. Log(1/*T*) (with $T = \text{energy}_{\text{bean}}/\text{energy}_{\text{reference}}$) measurements were collected at 2-nm increments over a range of 600−1898 nm. The time required for each bean, inclusive of loading, scanning, and cleanout, was ≈3 min. Following the scanning of each sample's beans, individual bean mass was measured to the nearest 0.1 mg.

Soybean meal was produced by grinding 25 g of each sample (≈20 s of grinding time) in a water-cooled mill (Knifetec 1095, Foss-Tecator, Hoganas, Sweden). The meal was used for measurement of P_i concentration and protein content by combustion, as well as NIR diffuse reflectance analysis. For the latter, ≈6 g of the grind was loaded into a standard forage ring cell for reflectance measurement. Diffuse reflectance [log(1/*R*), 1100−2498 nm, 2-nm increment, 32 scans per spectrum, 2 packs per sample, averaged] measurements were collected on the soybean meal samples. For both single-bean and meal spectra, data were exported from the Vision environment to external files of NSAS format. These files were read in and consolidated in the SAS macro program environment. All mathematical operations such as averaging were performed in SAS with the exception of the second-derivative transformation and the PLSR modeling, both of which were performed in Unscrambler (described below). Transfer of data into the Unscrambler worksheet was facilitated by the formation in SAS of the intermediate files in Excel format and eventual import of these files by Unscrambler.

Phosphorus and Protein Analysis. Inorganic phosphorus concentration was accomplished by colorimetry, as developed by Chen et al. (*13*) and modified by Raboy et al. (*1*), with additional modifications, described as follows. From the 30 g of seed that was ground in the sample mill described earlier, ≈100 mg was placed in a 1.50-mL microcentrifuge tube, to which 1 mL (10 μL/mg of seed tissue) of an extraction buffer (12.5% trichloracetic acid and 25 mM $MgCl_2$) was added. Tubes were vortexed and allowed to incubate for ≈16 h. Following additional vortexing and settling for 5 min, aliquots of extraction solution were transferred to 96-well storage plates and then centrifuged (3000 rpm) for 3 min. Extraction samples were measured in triplicate, with each subaliquot of 10 μL combined with 90 μL of ddH_2O and mixed with 100 μL of Chen's reagent (1 volume of 6 N H_2SO_4, 1 volume of 0.02 M ammonium molybdate, 1 volume of 10% ascorbic acid, and 2 volumes of water). Reaction time at ambient temperature was 1 h, whereupon P_i concentration was determined by a microplate spectrophotometer (BioTek Instruments, Winooski, VT), read at 882 nm, and calibrated beforehand with eight standards of P_i (Raboy's five, representing 0, 0.155, 0.465, 0.930, and 1.395 μg of P, plus three additional, 1.86, 2.32, and 2.64 μg of P, to extend the range).

Protein content (N × 6.25) was determined by combustion (*14*), using a Leco TruSpec CN (St. Joseph, MI) carbon−nitrogen analyzer. Single determinations of 150-mg portions of soybean meal were performed at laboratory ambient moisture conditions. The analyzer was calibrated with EDTA daily and checked with company-certified ground soybean meal.

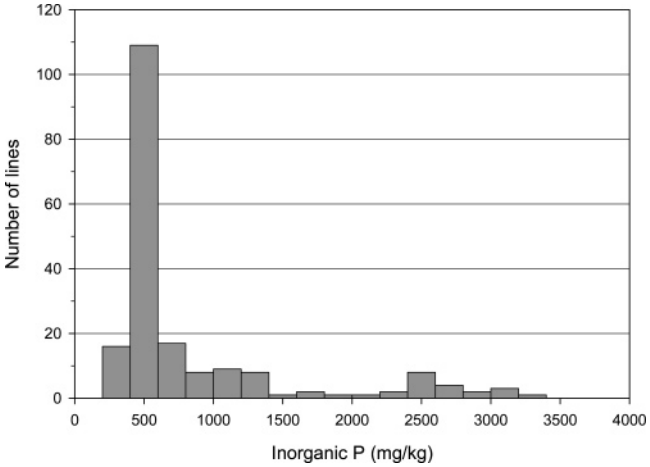

Figure 2. Histogram of inorganic phosphorus concentrations of samples.

PLSR Modeling. From the set of 191 samples, 150 were selected at random for calibration equation development; the remaining 41 samples were used for model validation. This structure was the same for soybean meal log(1/*R*) spectra and single-bean log(1/*T*) spectra. Spectra were pretreated with a Savitzky−Golay second derivative (15-point convolution window, quadratic polynomial) (*15*). The convolution window size of 15 points was selected on the basis of visual observation of the curves upon derivatization and preliminary regression modeling trials of widths ranging between 5 and 25 points (*16*), with 15 points demonstrating consistent, stable performance for both constituents and both spectral types. For the single-bean transmittance spectra, the lack of usable information in the low-wavelength region and noisy response in the high-wavelength region prompted the truncation of 100 and 50 points, respectively, thus reducing the usable region to 800−1798 nm. Transmittance spectra were processed in two forms: one as truly single-bean spectra, and the other as the average of 24 spectra per line. For both forms, reference values for protein content and P_i concentration were based on a portion from the 25-g grind of the sample.

PLSR was separately performed on each of the two constituents using commercial software (Unscrambler, v. 7.5, Camo AS, Trondheim, Norway). Full (one-sample-out rotation) cross-validation was used, such that the number of factors chosen for the regression equation generally corresponded to the minimum of the square root of the mean square of residuals (i.e., the rmsd) or one factor fewer. The saved regression equations were subsequently applied to the validation samples. Model accuracy was reported as the cross-validation coefficient of determination (R^2) and rmsd, as well as the validation set coefficient of determination (r^2) and rmsd. A dimensionless figure of merit, known as the RPD (*17*) and defined as the ratio of the validation set standard deviation of the constituent to the set's rmsd, was also determined.

Classification Modeling. Linear discriminant analysis (LDA) (*18*) was also used in a classification model in which samples below and above a P_i concentration of 500 mg/kg were assigned to separate classes, respectively. This boundary, as seen in the P_i distribution in **Figure 2**, corresponded to a nearly even split between samples into "low" and "high" categories. The same cross-validation and validation set structure as in the PLS regression analysis was used. Separate analyses were performed on the meal reflectance and 24-bean transmittance second-derivative spectral sets. For each set, the derivatized spectra were first reduced in dimension by principal component analysis (PCA), whereupon the four best scores for low versus high classification, as determined in stepwise analysis, were used as input to the LDA routine. All classification modeling was performed in SAS (*19*).

RESULTS

Descriptive statistics of the reference measurements of protein content and P_i concentration of the calibration and validation sets are listed in **Table 1**. The calibration set protein content ranged from 358 to 444 g/kg, with a mean of 395.2 g/kg and a standard deviation of 18.2 g/kg. The validation set protein content readings were quite similar. For P_i concentration, the calibration set range was 334−3370 mg/kg, with a mean and standard deviation of 836.1 and 772.6 mg/kg, respectively. In

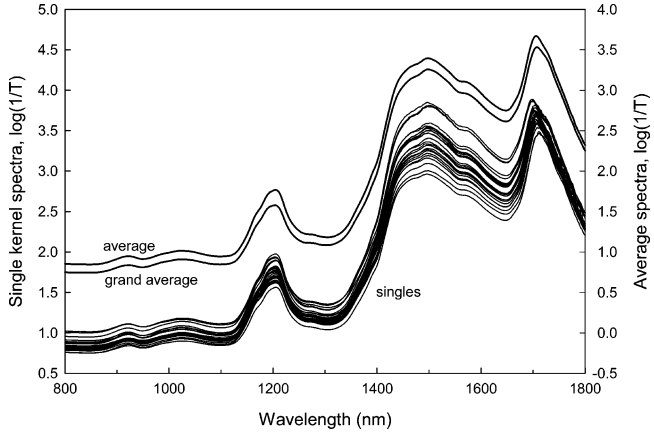

Figure 3. Typical transmittance spectra of 24 single soybeans of a sample line. The sample's average spectrum is also shown, as is the grand average spectrum that is based on all 191 samples (with 24 single-bean spectra per sample) used in the study.

Table 1. Descriptive Statistics of Soybean Lines[a]

constituent	set	*N*	mean ± SD	range
protein content (g/kg)	CV	150	395.2 ± 18.2	358–444
	validation	41	390.6 ± 17.4	360–425
inorganic P concn (mg/kg)	CV	150	836.1 ± 772.6	334–3370
	validation	41	839.4 ± 667.2	408–2610

[a] *N* = number of samples; CV = cross-validation.

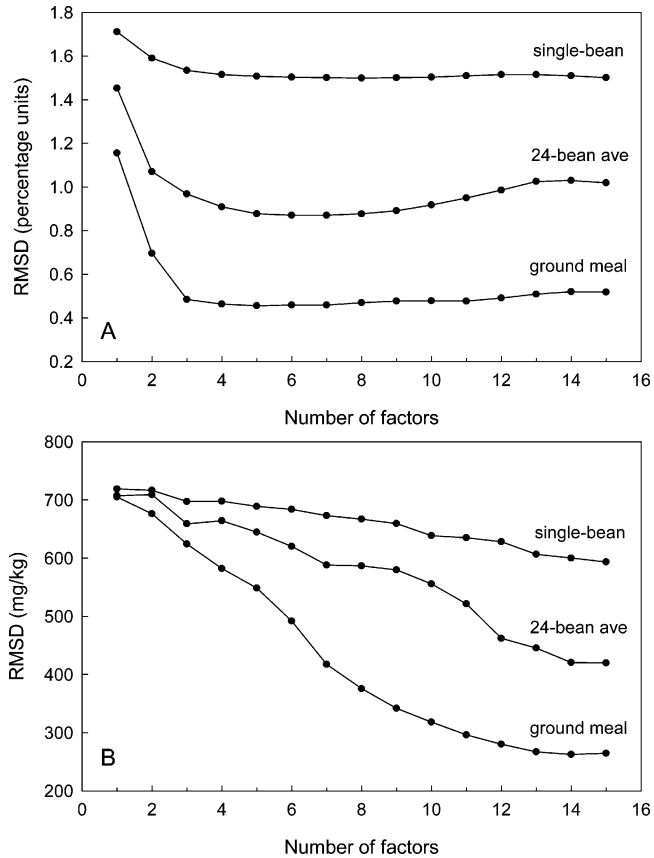

Figure 4. Cross-validation model errors for PLSRs based on the spectra of soybean meal, the average spectra of 24 soybeans per line, and single beans: (**A**) protein content; (**B**) inorganic phosphorus concentration.

Table 2. Summary of Partial Least-Squares Regression Model Results[a]

			cross-validation ($n = 150$)		validation ($n = 41$)		
constituent	spectral type	no. of factors	R^2	rmsd	r^2	rmsd	RPD
protein content (g/kg)	meal reflectance	5	0.937	4.55	0.931	4.64	3.8
	24-bean average transmittance	7	0.772	8.70	0.846	6.78	2.6
	single-bean transmittance	7	0.318	15.01	0.369	13.93	1.2
inorganic P concn (mg/kg)	meal reflectance	14	0.867	262.7	0.860	248.2	2.7
	24-bean average transmittance	12	0.592	461.9	0.589	444.5	1.5
	single-bean transmittance	13	0.294	606.4	0.276	568.6	1.2

[a] R^2 = multiple coefficient of determination of cross-validation samples; rmsd = square root of the mean of squared differences of residuals; r^2 = multiple coefficient of determination of validation samples; RPD = $SD_{validation}/rmsd_{validation}$.

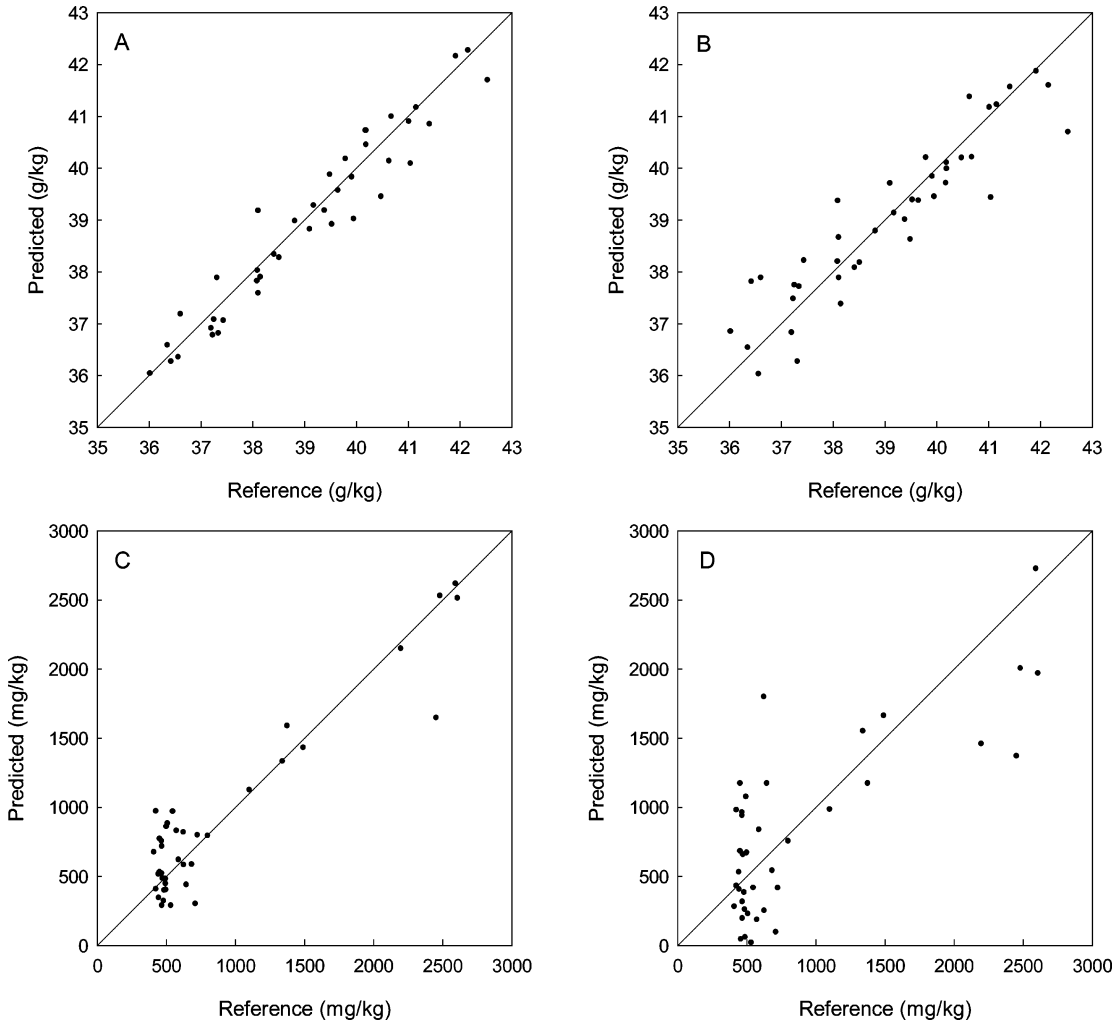

Figure 5. Predicted versus reference values for validation samples from PLS models of protein content (**A**, **B**) and inorganic phosphorus concentration (**C**, **D**). Graphs on the left side (**A**, **C**) correspond to soybean meal reflectance spectra, whereas graphs on the right side (**B**, **D**) correspond to the average of 24 single-bean transmittance spectra.

the umbrella study, a slight, although significant, correlation was observed between soil P and seed P_i ($r^2 = 0.22$) (*9*), which was supportive of the reported stronger correlation that exists between soil P and seed phytate (*20*). Again, the validation set statistics for P_i were correspondingly similar to the calibration set statistics. The distribution of inorganic phosphorus is depicted in **Figure 2**. Approximately half of the samples had P_i concentrations of <500 mg/kg. Some of the samples had P_i concentrations in excess of 2500 mg/kg, these being the samples that are expected to be of very low phytate content. Furthermore, all 14 samples in this high P_i category were homozygous for both phytate markers for the Cx1834-1-2 allele (*9*).

The spectra of a typical sample are presented in **Figure 3**. These spectra are characterized by four broad absorption peaks centered at 920, 1200, 1500, and 1700 nm, broadly attributed to protein C−H for the first wavelength and to C−H and O−H stretch of all three major compounds (protein, oil, carbohydrate) for the other wavelengths (*21*, *22*). Over the 800−1800 nm range, the largest variation occurred at the 1500 nm peak, where the range of spectral response was ≈0.8 log(1/*T*) units. In addition, the average spectrum of the 24 single-bean spectra is shown in the figure, as is the grand average spectrum, which is the average of all 4584 (= 24 × 191) single-bean spectra. Bean size, as measured by mass, was primarily responsible for the

spectral variation, as noted by a positive correlation ($r = 0.614$) between soybean mass and a spectral region of low nonspecific absorption, $\log(1/T)_{1650\ nm}$. With respect to the Teflon reference material, the level of transmitted energy within a spectrum varied by >2 orders of magnitude.

The cross-validation residual errors of PLSR equations for protein content and P_i are summarized in **Figure 4**, respectively. For protein content (graph **A**), five to seven PLS factors were deemed to be satisfactory. For P_i (graph B), a much larger number of PLS factors, 12−14, was needed. For both constituents, the models based on soybean meal spectra were substantially better than those based on either the average of 24 single-bean spectra or single-bean spectra alone. Using protein content as an example, the cross-validation rmsd values were 4.6, 8.7, and 15.0 g/kg for PLS equations that used soybean meal reflectance (five factors), the 24-bean average transmittance (seven factors), and the single-bean transmittance (seven factors), respectively. A similar trend occurred for the P_i models, as shown in **Table 2**. Considering that the single-bean models were reliant on reference chemical values determined from composites of hundreds of beans per sample, it is to be expected that the rmsd values for these models will be relatively large because of an inherent bean-to-bean variation of the constituent. Therefore, the fairer comparison of spectral measurement procedures involves that between the soybean meal reflectance and the 24-bean transmittance. In so doing, the rmsd of the 24-bean average protein model was found to be 91% larger than the corresponding soybean meal model; for P_i, it was 76% larger.

Validation statistics of the optimal PLSR equations for protein content and P_i are also contained in **Table 2**. Consistent with the cross-validation errors, validation set rmsd values were the smallest for soybean meal reflectance spectra (4.6 g/kg protein, 248 mg/kg P_i) and largest for single-bean transmittance spectra (13.9 g/kg protein, 569 mg/kg P_i). Expressed in nondimensional units, the RPD values for single-bean models were very low for both constituents (RPD = 1.2). The improvement in RPD value with spectral averaging was dependent on the analyte, as seen by the slight rise to 1.5 for P_i and the much larger rise to 2.6 for protein content. Because the single-bean RPD values approached 1.0, which indicates a lack of modeling power, it is doubtful that single-bean NIR models would be useful in breeding; however, this could change, provided the regression equations become based on reference readings performed on individual beans.

Plots of NIR-predicted versus reference measurements are shown for the validation samples in **Figure 5**. The plots correspond to the soybean meal reflectance and 24-bean average transmittance models for both protein content and P_i concentration. Protein contents were more uniformly distributed than P_i. The absolute value of the residual error for either modeled constituent was approximately constant throughout the constituent's range. Consistent with the rmsd values of **Table 2**, the 24-bean average transmittance models (**Figure 5B**,**D**) showed more dispersion from the 45-degree line than the soybean meal reflectance models (**Figure 5A,C**). The results of the P_i classification trials (P_i < 500 mg/kg vs P_i > 500 mg/kg) by LDA are contained in the confusion matrices of **Table 3**. On the basis of cross-validation, the accuracy of the meal reflectance model, with an average of 72.6% correct assignments, was slightly better than that of the 24-bean transmittance model (66.8%). However, both spectral types demonstrated nearly the same classification performance upon model validation (60.6 and 61.1% correctness averages for transmittance and reflectance types, respectively).

Table 3. Confusion Matrices for Classifications by Linear Discriminant Analysis[a]

			assigned category	
spectral type	set[b]	true P_i category	<500 mg/kg	>500 mg/kg
meal reflectance	CV	<500 mg/kg	51	21
		>500 mg/kg	20	58
	validation	<500 mg/kg	9	11
		>500 mg/kg	5	16
24-bean average transmittance	CV	<500 mg/kg	51	21
		>500 mg/kg	29	49
	validation	<500 mg/kg	13	7
		>500 mg/kg	9	12

[a] Input variables for linear discriminant analyses, according to spectral type, were as follows: meal reflectance, principal component (PC) score factors 5, 10, 2, and 8; 24-bean average transmittance, PC score factors 3, 9, 10, and 1. For each type, the spectral treatment prior to PC decomposition was a 15-point Savitzky–Golay quadratic polynomial second derivative. [b] CV = cross-validation ($n = 150$); validation ($n = 41$).

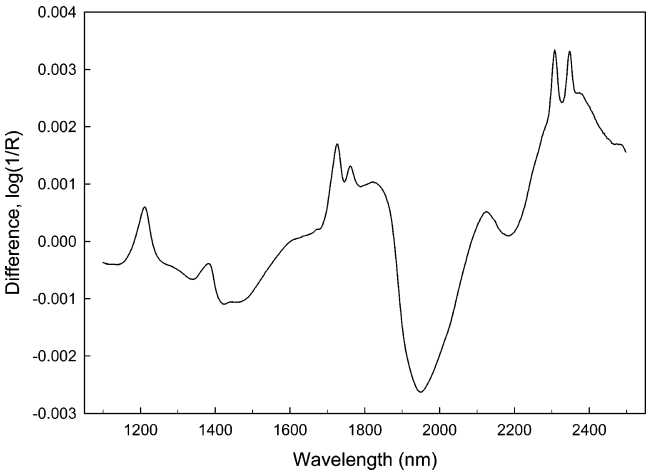

Figure 6. Difference spectrum, formed as the difference between the average spectra of the 30 highest and 30 lowest phosphorus calibration samples.

A difference spectrum that was formed as the average spectrum of the 30 highest P_i calibration set samples minus the average spectrum of the 30 lowest P_i calibration set samples is depicted in **Figure 6**. Sharp peaks occur at 1212, 1382, 1726, 1760, 2308, and 2348 nm, these being typically attributed to oil C−H stretch vibrations (*22*). It is reasoned that the greater absorption associated with these C−H bands in the high P_i samples arises from a smaller level of hydrogen bonding compared to the low P_i, hence, high-phytate samples. Conversely, a broad negative peak occurs in the 1900−2000 nm water O−H combination band region, which suggests that there also exists an inverse relationship between P_i and seed moisture.

DISCUSSION

The conventional procedure for assaying phosphorus in seed is by modification of the colorimetric method of Chen et al. (*13*). As such, seeds are ground or crushed, after which acid extraction and centrifugation are used to develop the phosphorus complex, a multihour procedure. To the plant breeder, the advantages of an NIR procedure are the savings in time, the simplicity of operation, and the ability to screen hundreds of lines daily.

Whereas the distribution of protein content was relatively uniform, the same is not true for P_i, as shown by **Figure 2**. This is often the case in breeders' sets in which the phenotypic expression is controlled by three or fewer genes. Although the relatively large deviation of the residuals of the P_i models may discourage their use in quantitative analysis, spectrally based classification models may be of sufficient accuracy to breeding programs. Often, the condition of the analyte being of high or low expression, indicative of the genotype, is of more importance to the breeder than the actual concentration. A closer examination of the classification assignments of the very high P_i samples, defined as $P_i > 1000$ mg/kg, revealed that 30 of 33 samples were correctly classified during cross-validation. For the corresponding validation set, 6 of the 8 very high P_i samples were correctly classified. These findings are a further suggestion of the enhanced usefulness of NIR methodology as a screening tool in soybean breeding.

In summary, NIR methodologies, inclusive of diffuse reflectance of soybean meal and, to a lesser extent, whole-bean transmittance, have the potential for use as tools in the screening of soybean lines for high inorganic phosphorus. This study demonstrates that NIR reflectance of soybean meal can be used to give a gross estimate of P_i concentration and as a means for classifying breeders' lines into categories of low and high P_i. Although single-bean NIR transmittance may also provide comparable information, additional research is needed at the level of phosphorus analysis of individual beans. Because the NIR methodology is in common use for the evaluation of oil and protein concentrations, breeding programs may benefit from the addition of phosphorus screening at little additional cost, thus demonstrating a low-cost alternative to the use of genetic markers or conventional wet chemical assay procedures.

ABBREVIATIONS USED

P_i, inorganic phosphorus; NIR, near-infrared; rmsd, square root of the mean squared differences between NIR-predicted and measured concentrations; PLSR, partial least-squares regression; LDA, linear discriminant analysis; PCA, principal component analysis.

SAFETY

The chemicals, equipment, and procedures of this study were handled and performed in accordance with usual precautionary measures.

ACKNOWLEDGMENT

We thank B. Stetzler and S. Mookherji (ARS, Beltsville) for spectral data collection and combustion protein analyses.

LITERATURE CITED

(1) Raboy, V.; Gerbasi, P. F.; Young, K. A.; Stoneberg, S. D.; Pickett, S. G.; Bauman, A. T.; Murthy, P. P. N.; Sheridan, W. F.; Ertl, D. S. Origin and seed phenotype of maize low phytic acid 1-1 and low phytic acid 2-1. *Plant Physiol.* **2000**, *124*, 355−368.

(2) Oltmans, S. E.; Fehr, W. R.; Welke, G. A.; Raboy, V.; Peterson, K. L. Agronomic and seed traits of soybean lines with low-phytate phosphorus. *Crop Sci.* **2005**, *45*, 593−598.

(3) Raboy, V.; Dickinson, D. B.; Below, F. E. Variation of seed total phosphorus, phytic acid, zinc, calcium, magnesium, and protein among lines of *Glycine max* and *G. soja*. *Crop Sci.* **1984**, *24*, 431−434.

(4) Sharpley, A. N.; Charpa, S. C.; Wedepohl, R.; Sims, J. Y.; Daniel, T. C.; Reddy, K. R. Managing agricultural phosphorus for protection of surface waters: issues and options. *J. Environ. Qual.* **1994**, *23*, 437−451.

(5) Wilcox, J. R.; Premachandra, G. S.; Young, K. A.; Raboy, V. Isolation of high seed inorganic P, low-phytate soybean mutant. *Crop Sci.* **2000**, *40*, 1601−1605.

(6) Hitz, W. D.; Carlson, T. J.; Kerr, P. S.; Sebastian, S. A. Biochemical and molecular characterization of a mutation that confers a decreased raffinosaccharide and phytic acid phenotype on soybean seeds. *Plant Physiol.* **2002**, *128*, 650−660.

(7) Oltmans, S. E.; Fehr, W. R.; Welke, G. A.; Cianzio, S. R. Inheritance of low-phytate phosphorus in soybean. *Crop Sci.* **2004**, *44*, 433−435.

(8) Walker, D. R.; Scaboo, A. M.; Pantalone, V. R.; Wilcox, J. R.; Berma, H. R. Genetic mapping of loci associated with seed phytic acid content in Cx1834-1-2 soybean. *Crop Sci.* **2006**, *46*, 390−397.

(9) Scaboo, A. M. Validation of quantitative trait loci conditioning seed phytate and protein concentration in soybean [*Glycine max* (L.) Merrill]. M.S. thesis, University of Tennessee, Knoxville, TN, 2005.

(10) Raboy, V. Progress in breeding low phytate crops. *J. Nutr.* **2002**, *32*, 503S−507S.

(11) Pantalone, V. R.; Allen, F. L.; Landau-Ellis, D. Registration of '5601T' soybean. *Crop Sci.* **2003**, *43*, 1123−1124.

(12) Fehr, W. R. *Principles of Cultivar Development, Vol. I: Theory and Techniques*; Iowa State University Press: Ames, IA, 1991.

(13) Chen, P. S., Jr.; Toribara, T. Y.; Warner, H. Microdetermination of phosphorus. *Anal. Chem.* **1956**, *28*, 1756−1758.

(14) AACC International. *Crude Protein—Combustion Method*; AACC Method 46-30; American Association of Cereal Chemists: St. Paul, MN, 2000.

(15) Savitzky, A.; Golay, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1963**, *36*, 1627−1639.

(16) Reeves, J. B., III; Delwiche, S. R. SAS partial least squares regression for analysis of spectroscopic data. *J. Near-Infrared Spectrosc.* **2003**, *11*, 415−431.

(17) Williams, P. C., Implementation of near-infrared reflectance technology. In *Near-Infrared Technology in the Agricultural and Food Industries*, 2nd ed.; Williams, P. C., Norris, K. H., Eds.; American Association of Cereal Chemists: St. Paul, MN, 2001; pp 145−169.

(18) Fisher, R. A. The use of multiple measurements in taxonomic problems. *Ann. Eugenics* **1936**, *7*, 179−188.

(19) SAS Institute Inc. The DISCRIM, PRINCOMP, and STEPDISC Procedures. In *SAS/STAT User's Guide*, release 6.03 ed.; SAS Institute: Cary, NC, 1988; pp 359−447, 751−771, 909−922.

(20) Raboy, V.; Dickinson, D. B. Phytic acid levels in seeds of *Glycine max* and *G. soja* as influenced by phosphorus status. *Crop Sci.* **1993**, *33*, 1300−1305.

(21) Osborne, B. G.; Fearn, T. *Near Infrared Spectroscopy in Food Analysis*; Longman: Essex, U.K., 1986; pp 20−42.

(22) Murray, I.; Williams, P. C. Chemical principles of near-infrared technology. In *Near-Infrared Technology in the Agricultural and Food Industries*; Williams, P. C., Norris, K. H., Eds.; American Association of Cereal Chemists: St. Paul, MN, 1987; pp 17−34.

Received for review March 31, 2006. Revised manuscript received July 19, 2006. Accepted July 24, 2006. Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply endorsement or recommendation by the USDA or the University of Tennessee.

JF060919N